

Pred. No. is the number of results predicted by chance to have a

Compositions and methods for the detection and identification of drugs, poisons, and explosives.

JOURNAL Patent: WO 0216581-A 3 28-FEB-2002;
Genentech, Inc. (US)

FEATURES Location/Qualifiers
1.3437

SOURCE /organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 558 a 1186 c 1098 g 595 t
ORIGIN

Query Match Best Local Similarity 100.0%; Score 3437; DB 6; Length 3437;
Matches 3437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGACCAAGGCTTCTTCTACCTGAGAGCGGGGAGAGACCAACCAATCATCTCTG 60
DB 1 CAGGACCAAGGCTTCTTCTACCTGAGAGCGGGGAGAGACCAACCAATCATCTCTG 60
QY 61 GTCCATGCCATGATGATCTCTGATGAGCGCTGGCCCGCTCTGAGCCGAGACAGGATTC 120
DB 61 GTCCATGCCATGATGATCTCTGATGAGCGCTGGCCCGCTCTGAGCCGAGACAGGATTC 120
QY 121 CAGGCGCTGCTGAGCATCTGCTTTCCGAGAGAGAAAGCACTGCCACCGCTTCTGCTG 180
DB 121 CAGGCGCTGCTGAGCATCTGCTTTCCGAGAGAGAAAGCACTGCCACCGCTTCTGCTG 180
QY 181 GACACATGAGAGAGAGCGGTGCTCTCTGACCTGAGCTGAGCTGAGCTGAGCTGCT 240
DB 181 GACACATGAGAGAGAGCGGTGCTCTCTGACCTGAGCTGAGCTGAGCTGAGCTGCT 240
QY 241 GAGGTGCTCCCGCTGCTGAGAGCGCGCTGAGAGACCTGAGAGACAGCTGCTGCTG 300
DB 241 GAGGTGCTCCCGCTGCTGAGAGCGCGCTGAGAGACCTGAGAGACAGCTGCTGCTG 300
QY 301 TTCTGCAAGTCTGTTGGCATCCCGCTGTCAGAGATAGAAATCTCTCACTTCTGAGAC 360
DB 301 TTCTGCAAGTCTGTTGGCATCCCGCTGTCAGAGATAGAAATCTCTCACTTCTGAGAC 360
QY 361 CAGGAGTGGCCACAGACCCCACTCTGAGACAGAAATCAATGAGAGCAAGATTAATG 420
DB 361 CAGGAGTGGCCACAGACCCCACTCTGAGACAGAAATCAATGAGAGCAAGATTAATG 420
QY 421 GCGCACTGGTGGAGGTCCAGATAGAGCGCGCTCCGAGAGCCAGACTTCTCACTCC 480
DB 421 GCGCACTGGTGGAGGTCCAGATAGAGCGCGCTCCGAGAGCCAGACTTCTCACTCC 480
QY 481 TTGCTCAAGCTCTCCCTGCGCGCGCGAGAGACAGAAAGAGATCCCAAAAGAGC 540
DB 481 TTGCTCAAGCTCTCCCTGCGCGCGCGAGAGACAGAAAGAGATCCCAAAAGAGC 540
QY 541 AGCCCAAGAGACCCATAGGCGAGGCGGATTCGGGTGGGAGCCAGCTCCGGGTCTG 600
DB 541 AGCCCAAGAGACCCATAGGCGAGGCGGATTCGGGTGGGAGCCAGCTCCGGGTCTG 600
QY 601 GAGCCCTAGAGACCACTGCTGAGTATCTCTCAAGATTTTCCCGCTACCCCGAGACCT 660
DB 601 GAGCCCTAGAGACCACTGCTGAGTATCTCTCAAGATTTTCCCGCTACCCCGAGACCT 660
QY 661 CGGTGGAGAGTCCAGTCCCGCGCGCGAGGCGCTCCGAGAGAGCCCTGGGCGAG 720
DB 661 CGGTGGAGAGTCCAGTCCCGCGCGCGAGGCGCTCCGAGAGAGCCCTGGGCGAG 720
QY 721 GAGTGGCCCGGTGCTCAAGGAGAGCCCGAGGTCCGAGGACATAGTGTGCTGCTG 780
DB 721 GAGTGGCCCGGTGCTCAAGGAGAGCCCGAGGTCCGAGGACATAGTGTGCTGCTG 780
QY 781 CAGGCGCTGCGACAGCTGCTCACTCCCAACAGCGCGCTGCTGCTGATGCTCATCAG 840
DB 781 CAGGCGCTGCGACAGCTGCTCACTCCCAACAGCGCGCTGCTGCTGATGCTCATCAG 840
QY 841 CGTAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 841 CGTAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

QY 901 CCACAGACACCGGCTTCTCTGCTCTTCTGAGGTCTCTGAGAGGTCTGAGAGG 960
DB 901 CCACAGACACCGGCTTCTCTGCTCTTCTGAGGTCTCTGAGAGGTCTGAGAGG 960
QY 961 CTGAGACACCGGCTTCTGAGAGGCGGGCCCTGCGGACACAGTCAAGATGCTTCCAGC 1020
DB 961 CTGAGACACCGGCTTCTGAGAGGCGGGCCCTGCGGACACAGTCAAGATGCTTCCAGC 1020
QY 1021 CAGGCGCTGAGAGGCGGAGAGGTGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGG 1080
DB 1021 CAGGCGCTGAGAGGCGGAGAGGTGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGG 1080
QY 1081 GCGCTGAGCTTCCCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGT 1140
DB 1081 GCGCTGAGCTTCCCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGT 1140
QY 1141 ACCCTGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAG 1200
DB 1141 ACCCTGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAG 1200
QY 1201 GAGGCGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGG 1260
DB 1201 GAGGCGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGG 1260
QY 1261 ACTGCGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGG 1320
DB 1261 ACTGCGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGGAGTCTGAGAGG 1320
QY 1321 CTGCTGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1321 CTGCTGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
QY 1381 GAGGCGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
DB 1381 GAGGCGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
QY 1441 CCGAGAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
DB 1441 CCGAGAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
QY 1501 GCGAAAGGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
DB 1501 GCGAAAGGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
QY 1561 CAGTCCAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
DB 1561 CAGTCCAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
QY 1621 CAGAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
DB 1621 CAGAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
QY 1681 ATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
DB 1681 ATCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
QY 1741 GTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
DB 1741 GTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1800
QY 1801 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
DB 1801 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
QY 1861 AGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
DB 1861 AGTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
QY 1921 ATCCAGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
DB 1921 ATCCAGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
QY 1981 TACCTACAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040

Db	3061	ATGTGATGGCGCATGATGACCCACAGCGCCAGATGTCGAGGCGCCCTGAGAGATCCTCATATG	3120
QY	3121	GAGGCGCGGATGATGAGCCTCTGTGTGTGCAAGCCACCCCTCTCCAGGCCCGCGGCGCGTCCGCGTCC	3180
Db	3121	GAGGCGCGGATGATGAGCCTCTGTGTGTGCAAGCCACCCCTCTCCAGGCCCGCGGCGCGTCCGCGTCC	3180
QY	3181	CGGAGGATCTCTCGAGGCGCAAGGCCCAAGGAAAGCTGGGCGTGTCTGTCTGTCTGTCCGAGAGGT	3240
Db	3181	CGGAGGATCTCTCGAGGCGCAAGGCCCAAGGAAAGCTGGGCGTGTCTGTCTGTCTGTCCGAGAGGT	3240
QY	3241	GAGGCGCGGAGCCCTGAGGCGCAGGCGCAGGCGCCAGGAGCATATCTCGAGGCCCTGGGGTGG	3300
Db	3241	GAGGCGCGGAGCCCTGAGGCGCAGGCGCAGGCGCCAGGAGCATATCTCGAGGCCCTGGGGTGG	3300
QY	3301	CTTCGCGGCGCGGCGCGTGGCATCAAGGGCGCGTTCAGCAGAACCCCTATTACCTCTGTGGGCG	3360
Db	3301	CTTCGCGGCGCGGCGCGTGGCATCAAGGGCGCGTTCAGCAGAACCCCTATTACCTCTGTGGGCG	3360
QY	3361	ACAGGCCCTGCCCGGGAAGCGGGGAGTCCCGCGGGCATAGGCGTGGGCTGTTTGAATGAA	3420
Db	3361	ACAGGCCCTGCCCGGGAAGCGGGGAGTCCCGCGGGCATAGGCGTGGGCTGTTTGAATGAA	3420
QY	3421	ACGACCTGGAACCTGTCAA	3437
Db	3421	ACGACCTGGAACCTGTCAA	3437

RESULT 2				
AB037861	AB037861	4434 bp	mRNA	linear
LOCUS				PRI 14-MAR-2000
DEFINITION	Homo sapiens mRNA for KIAA1440 protein, partial cds.			
ACCESSION	AB037861			
VERSION	AB037861.1	GI:7243277		
KEYWORDS				
SOURCE	Homo sapiens brain cDNA to mRNA, clone_1ib:pluescriptII SK plus			
ORGANISM	Homo sapiens			
	Clone:hg02441b.			

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE REFERENCE AUTHORS TITLE
Nagase, T., Kikuno, R., Ishikawa, K.-I., Hirosewa, M. and Ohara, O. Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro	DNA Res. 7 (1), 65-73 (2000) 20181126
2 (bases 1 to 4434) Ohara, O., Nagase, T. and Kikuno, R. Direct Submission	
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: odnainfo@kazusa.or.jp, URL: http://www.kazusa.or.jp/huge/ ; Tel: +81-438-52-3913, Fax: +81-438-52-3914)	

FEATURES

source	Location/Qualifiers
gene	1..4434
CDS	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="hg02441b" /tissue_type="brain" /clone_lib="pblscriptlit SK plus" 1..4434 /gene="KIAL1440" <1..4134 /gene="KIAL1440" /note="Start codon is not identified." /codon_start=1 /product="KIAL1440 protein" /protein_id="BA92678.1" /db_xref="gi:7243278" /translation="AASSTKQTIESSSLTSLSLDPGGRRPPPHILDYKSLN OSIRGHILCRSNPFLILTIORASQSNMADLYOSSSGSLDVPVQICCEFL HAYVDAASEEDDESGEKAKKRRQKROKRIILPLPKADSRQTCFELD

YFRLGSSQVASYNLAMKGLSLVSEGLRDGEKBPWEEDVDTLVQYOMLIR
 DFLRLDVSASTTALALQALHMETDPTSAVLYSSORPPEBOADSLADYD
 AKLVERKSIHSLFKISPSASDAVETLSTFYSYRMRORKEEYPSMSQ
 DQYELMSGEGATMLIVYAVITLLDPPRDESEHQALDITWPEEYPSMSQ
 VDTSEALLDPMILRLIBSEYLRVYALDIPDQALITVOSFGLVPSMSKLO
 ELDOAYADPOTLQNTMKNIMAHVEOHERGASGGTFFSLTALSPRDETEA
 PAKSSPDPITGQGRIVGTQIVLQALITLSPRDETEA
 LQALQELALAVYQSPVGLIVYVQALITLSPRDETEA
 QLOQYRCVDPDGTFSFLKYLQMLDMLSPRDETEA
 DVGGLRLALALAVYQSPVGLIVYVQALITLSPRDETEA
 PHEELTAFSAADASPACRYVYSSLLDSEEPVAGKRPADGSSAEVRL
 GPSSGLVDMILMDPVYVSCPDQILRLSRKRGQAVPSRYLLTFLTHQS
 WPLHOCIRVLGKSRKRDPSADILMACIHYRIMQGDQFPTOKREELVYLA
 QOPELISVELILAEATRSODDPAACSLIHYRIMQGDQFPTOKREELVYLA
 CLOMDSEVILGRCDLLOLYORPELRVYVYVSEGAASSYCKDILHRT
 TILADTSDRLNENRGDSMACRTLANPRLILHMLTALHGHGTLNQEORQ
 ONHSCPLVAVILGILHLOPHYTRSHQALMDLISFLILANTKRSRMLAFIKF
 VOFIRKITYTAPALISFLQKADLHDLSPDSDMLSLGSLISLPRDRTRG
 LDBEBSBSAGSLPLVSVLFTPLTAAEAPYKMLSGQVDELEVSLDIDEMSR
 RREELISFSTNLQRLMSSAEBCRNLAFLSLRMONSPSTIAAFIPMYCTGSD
 FEVYOTALRNLPEYALICOEHAVALIRALVGMQDPSRQISELRILHEAVY*

BASE COUNT 731 a 1521 c 1419 g 763 t

Query Match 99.8%, Score 3431.8; DB 9; Length 4434;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAGGACGAGGCTTCTCTAGCGGAGAGAGGAGAGAGGACCAACATGACATCTCTG 60
 DB 1000 CAGGACGAGGCTTCTCTAGCGGAGAGAGGAGAGAGGACCAACATGACATCTCTG 60
 QY 61 GTCCATGCAATGATGATCTCTGACGCTGGGCGGCTCTGAGCGGAGAGAGAGATC 120
 DB 1060 GTCCATGCAATGATGATCTCTGAGCGGAGAGAGAGAGAGATC 1119
 QY 121 CAGGCGTCTGAGACATCTGTTCCGAGAGAGAGAGAGAGAGAGAGAGAGATC 180
 DB 1120 CAGGCGTCTGAGACATCTGTTCCGAGAGAGAGAGAGAGAGAGAGAGAGATC 1179
 QY 181 GACACATCGATC 240
 DB 1180 GACACATCGATC 1239
 QY 241 GAGGTGCTGCGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 300
 DB 1240 GAGGTGCTGCGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1299
 QY 301 TTGCTGCAAGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 360
 DB 1300 TTGCTGCAAGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1359
 QY 361 CAGGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 420
 DB 1360 CAGGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1419
 QY 421 GCGGACCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 480
 DB 1420 GCGGACCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1479
 QY 481 TTGCTGCAAGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 540
 DB 1480 TTGCTGCAAGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1539
 QY 541 AGCCGATC 600
 DB 1540 AGCCGATC 1599
 QY 601 GCGGCTGATC 660
 DB 1600 GCGGCTGATC 1659
 QY 661 CCGTGGATC 720

DB 1660 CCGTGGATC 1719
 QY 721 GAGGTGCGCGGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 780
 DB 1720 GAGGTGCGCGGCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1779
 QY 781 CAGGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 840
 DB 1780 CAGGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1839
 QY 841 CGTGGACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB 1840 CGTGGACACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1899
 QY 901 CCAAGATC 960
 DB 1900 CCAAGATC 1959
 QY 961 CTGGATC 1020
 DB 1960 CTGGATC 2019
 QY 1021 CAGGCTGACCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1080
 DB 2020 CAGGCTGACCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2079
 QY 1081 GCGCTGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1140
 DB 2080 GCGCTGCGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2139
 QY 1141 ACCCTGATC 1200
 DB 2140 ACCCTGATC 2199
 QY 1201 GCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1260
 DB 2200 GCGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2259
 QY 1261 ACTGGATC 1320
 DB 2260 ACTGGATC 2319
 QY 1321 CTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1380
 DB 2320 CTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2379
 QY 1381 GAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1440
 DB 2380 GAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2439
 QY 1441 CCGGATC 1500
 DB 2440 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2499
 QY 1501 GCGCAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1560
 DB 2500 GCGCAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2559
 QY 1561 CAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1620
 DB 2560 CAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2619
 QY 1621 CAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1680
 DB 2620 CAGAGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2679
 QY 1681 ATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1740
 DB 2680 ATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 2739
 QY 1741 GTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1800

QY	2881	GGCCTGCGGTCCATGCAAGAAACAGCACCCTCAGCATTTGCAAGCCGCTTCTTCGCGCCACGTCATG	2940
Db	3880	GGCCTCGCTCCATGCAAGAAACAGCACCCTCAGCATTTGCAAGCCGCTTCTTCGCGCCACGTCATG	3939
QY	2941	TACTGCTGGGCACGCCAGCACTTTGAGGTGTGGCAAGGGCTCTCCGGAACTTCGCTGAG	3000
Db	3940	TACTGCTGGGCACGCCAGCACTTTGAGGTGTGGCAAGGGCTCTCCGGAACTTCGCTGAG	3999
QY	3001	TACGCTCTCTGTGCCAAGAGACGCGGCTGTGTCTCCACGGAGCTTCCTGTGGGGC	3060
Db	4000	TACGCTCTCTGTGCCAAGAGACGCGGCTGTGTCTCCACGGAGCTTCCTGTGGGGC	4059
QY	3061	ATGTTAGGCGCGAGTGGAGACCCACGGCCCAAGTCTCGAGGCCCTGAGAGATCCGATAG	3120
Db	4060	ATTTATAGGCGCGAGTGGAGACCCACGGCCCAAGTCTCGAGGCCCTGAGAGATCCGATAG	4119
QY	3121	GAGGCGCGTGAATGTAGAGCCGTGGGGAGGCCGACCCCTTCGAAAGCCCGGCGCTCCGCTCC	3180
Db	4120	GAGGCGCGTGAATGTAGAGCCGTGGGGAGGCCGACCCCTTCGAAAGCCCGGCGCTCCGCTCC	4179
QY	3181	CGGGGAGATCTCTCAGAGCAAAAGCCGAGAGAGCTGTGAGCTGTGCTGTGCCAGAGAGGT	3240
Db	4180	CGGGGAGATCTCTCAGAGCAAAAGCCGAGAGAGCTGTGAGCTGTGCTGTGCCAGAGAGGT	4239
QY	3241	GAGGGGCGCCAGAGCCCTGAGGCCAGAGCGAGGCCCAAGGACCAATACTCCGAGCCCTGGGGTGG	3300
Db	4240	GAGGGGCGCCAGAGCCCTGAGGCCAGAGCGAGGCCCAAGGACCAATACTCCGAGCCCTGGGGTGG	4299
QY	3301	CTTCGCGGCGCGGCGCGCTGGGAGATCAAGGGGGCGGTCCAGAGCAAGCCTCACTTACCTTCTGGGCC	3360
Db	4300	CTTCGCGGCGCGGCGCGCTGGGAGATCAAGGGGGCGGTCCAGAGCAAGCCTCACTTACCTTCTGGGCC	4359
QY	3361	AACAGCCCTGCCCGGAGAGGGGGAGTCCGCCCGGGGCAAGGCTGTGGGCTGTGTTGAATGGA	3420
Db	4360	AACAGCCCTGCCCGGAGAGGGGGAGTCCGCCCGGGGCAAGGCTGTGGGCTGTGTTGAATGGA	4419
QY	3421	AACGACCTCGAACCTGTG	3435
Db	4420	AACGACCTCGAACCTGTG	4434

BC004286 4179 bp mRNA linear PRI 13-JUL-2001
Homo sapiens, clone IMAGE:3618123, mRNA, partial cds.
BC004286
BC004286.1 GI:13279124
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumelostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4179)
Strausberg,R.
Direct Submission
Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCDP/DMP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Jetticia Hsiao, Martin Krywinski, Reta Kutsche, Oliver Lee, Soo
Sun Lee, Victor Lang, Carrie Mathewson, Candice McLeavy, Steven

QY	2641	TCACGCGGCGCTTTCGCCCCGCGTGAAGCGTCTCCCTCTTCAACCCTCTGACGCGGCGGAG	2700
Db	3353	TCACGCGGCTCTCTTCGCCCCGCGTGAAGCGTCTCCCTCTTCAACCCTCTGACGCGGCGGAG	3412
QY	2701	ATGGCCCTCCCTACATGAAGAGCGGCTTTCGCGGGGGCCAAAGCGTGAAGCATCGCTGGAAGTT	2760
Db	3413	ATGGCCCTCCCTACATGAAGAGCGGCTTTCGCGGGGGCCAAAGCGTGAAGCATCGCTGGAAGTT	3472
QY	2761	CTGAGTGAATATGACGAGATGTCCCGCGGAGAACCCGAGATCTGAGCTTCTTCGACCC	2820
Db	3473	CTGAGTGAATATGACGAGATGTCCCGCGGAGAACCCGAGATCTGAGCTTCTTCGACCC	3532
QY	2821	AACCTGACCGCGCGATGAGAGCTCGCGGCCAAGAGATGTTCGCGCAACTCGCTTCAGGCTG	2880
Db	3533	AACCTGACCGCGCGATGAGAGCTCGCGGCCAAGAGATGTTCGCGCAACTCGCTTCAGGCTG	3592
QY	2881	GCCCTCGCTCCATATGACGAGACCGCCAGATTTGCAGGCGGTTTCGTCGCGCAGCTGTATAG	2940
Db	3593	GCCCTCGCTCCATATGACGAGACCGCCAGATTTGCAGGCGGTTTCGTCGCGCAGCTGTATAG	3652
QY	2941	TACTGCTCTGGGCAACCAAGACATTTGATGTGTGTGCAAGACGGCCCTCCGAGACCTGCTGAG	3000
Db	3653	TACTGCTCTGGGCAACCAAGACATTTGATGTGTGTGCAAGACGGCCCTCCGAGACCTGCTGAG	3712
QY	3001	TACGCTCTCTGTCCTCCAAAGAGAGCGCGGCTGTGTGCTCCACCGGCGCTTCTGTGTGTGAGC	3060
Db	3713	TACGCTCTCTGTCCTCCAAAGAGAGCGCGGCTGTGTGCTCCACCGGCGCTTCTGTGTGTGAGC	3772
QY	3061	ATTATAGGCGCATATGAGACCCCAAGCGCGCATATCTCCAGAGCCCTTGAGGATCTGTGATAG	3120
Db	3773	ATTATAGGCGCATATGAGACCCCAAGCGCGCATATCTCCAGAGCCCTTGAGGATCTGTGATAG	3832
QY	3121	GAGGCGCGATGTGTGAGACCTGT	3180
Db	3833	GAGGCGCGATGTGTGAGACCTGT	3892
QY	3181	CGGCGGATCTTCGAGGCGCAAAAGCCCAAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3240
Db	3893	CGGCGGATCTTCGAGGCGCAAAAGCCCAAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3952
QY	3241	GAGGCGCGGAGACCCCTTGAGGCGGAGCGCGGCGGAGCAATATCTCGAGCCCTGTGTGTGT	3300
Db	3953	GAGGCGCGGAGACCCCTTGAGGCGGAGCGCGGAGCAATATCTCGAGCCCTGTGTGTGTGT	4012
QY	3301	CTTCGCGGCGCGGCGGCTGT	3360
Db	4013	CTTCGCGGCGCGGCGGCTGT	4072
QY	3361	ACAGCCCTTCGCGGAGAGCGGCGATCTCCCGCGGCGCATGAGCTGTGTGTGTGTGTGTGTGT	3420
Db	4073	ACAGCCCTTCGCGGAGAGCGGCGATCTCCCGCGGCGCATGAGCTGTGTGTGTGTGTGTGTGT	4132
QY	3421	ACGACCTGGAACCTGTGATA	3437
Db	4133	ACGACCTGGAACCTGTGATA	4149

RESULT 4					
EC013367					
LOCUS	BC013367	2769 bp	mRNA	linear	PRI 04-SEP-2001
DEFINITION	Homo sapiens, clone IMAGE:3677373,		mRNA,	partial cds.	
ACCESSION	BC013367				
VERSION	BC013367.1	GI:15426515			
KEYWORDS					
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 2769)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (31-AUG-2001) National Institutes of Health, Mammalian				

Db	2409	GAGGATCTCTGCATATGAGAGGCCCGTGAATGTGTAGCTGTGTGGCAGGCCACCCCTCTCAAGCC	2458
QY	3165	CCGGGCGCCCTCCGCTCCCGGGGGATCCGCCAGGCAAAAGCCAGAAAGCGTGGCGTGGCTG	3224
Db	2469	CCGGCCCGCTCCCTCCCTCCCGGGGAATCTCCAGGCAAAAGCCAGAAAGCGTGGCGTGGCTG	2528
QY	3225	GTCTGTCCGAGGAGGTGAGGGCGCCAGGCCCTGAGGCGCAGGCAAGCCAGAGCAATAT	3284
Db	2529	GTCTGTCCGAGGAGGTGAGGCGCGCCAGGCCCTGAGGCAAGCCAGAGCAATAT	2588
QY	3285	CCGAGCCCTGGGGTGGCTCCGGGTCGGGCGGCTGGCATCAAGGGGCGTCCAGCAAGCCCTC	3344
Db	2589	CCGAGCCCTGGGGTGGCTCCGGGTCGGGCGGCTGGCATCAAGGGGCGTCCAGCAAGCCCTC	2648
QY	3345	ATTTCACCTTCTGGGSCCACAAGCCCTGGCCGCGAGAGGGGAGATCTCCCCC-GAGCATGGGCTG	3403
Db	2649	ATTTCACCTTCTGGGSCCACAAGCCCTGGCCGCGAGAGGGGAGATCTCCCCC-GAGCATGGGCTG	2708
QY	3404	GCGTGGTTTGGAATGAACGACGCTGSAACGTGCAC	3437
Db	2709	GCGTGGTTTGGAATGAACGACGCTGSAACGTGCAC	2742

1

2418 DP, mRNA linear PRI 11-DEC-2001
Homo sapiens, clone IMAGE:486992, mRNA, partial cds.
BC018777
BC018777.1 GI:17511852
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2418)
Strausberg,R.
Direct Submission
Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIG-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DTP
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Genotyping by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Kravynski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhun, Parvaneh Saeedi, Jacqueline
Schain, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorpe, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
George Yang, Scott Zuydamdyun, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 40 Row: K Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers
1. 2418
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:486992"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIG_MGC_49"
/lab_host="DH10B-R"

CDS

/note="Vector: pOTB7"
<1..2097
/codon_start=1
/product="Unknown (protein for IMAGE:469992)"
/protein_id="AAH1877.1"
/db_xref="GI:17511853"

/translation="RLSDVKGGLILALBALAFRODLEVVSTVRAYVATLRSQCSY
EPDLISKVLOGLIEVSPHLEILAFPSAFAADPAPACPVVYSILLDEEPL
AGKPGADGGSLIENVRAGPSGLIVDLEMDEPVSQDLOLILFSTRKKGQAO
VSEFPEYLLTLETHOSMPLTHOCIRVILICKRSQEPDSASLDFWACIHVRLWQ
RODRTPOKREELVRYOPELISVELLAEVTSQDQAPCSLQARPELLSC
CGDDESTRKATYHELSCQIDQDVSILGRKQDLQILOPELRYVPEVILHSG
AASVSVKLDGLIHRFTLLADTSDSRLNRRDASNAKRLAVARLILHLEPMI
AALLKRLHLEFQEFQONLSCEFLQYHLELOPHVRSHOGALMDLLSFRIL
LYKRSRHLAEFINKVOYFIKTYTVAAPALISLOKAPIDHLSFPMDSIDVILKS
LLAGSLPSRDDRDGLDEEESSSAGSLPYVSYLEPILAEAPYMKRLSLWS
TYEDILEYSDIDMSRRRPELISFTSTINQIRLMSAECCNIAFSLARSONPS
IAAFLPTMYCIGSDQFEVVTALRNIPYALLCOEHAVALHRAVLVGMQMDPS
AOLISALILHMERAVM"

BASE COUNT 411 a 811 c 776 g 420 t
ORIGIN

Query Match 69.8%; Score 2397.4; DB 9; Length 2418;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1039 AGGCTCAGTATGTGAGAGGGGGCTCTGCGCTGCGCGAGGCCCTTCGCTAG 1098
DB 1 AGGCTCAGTATGTGAGAGGGGGCTCTGCGCTGCGCGAGGCCCTTCGCTAG 60
QY 1099 GAGCTGAGAGTGTGAGCTCCAGCTCCGCTCCGCTCATGCGCACTGAGTCTGGGAG 1158
DB 61 GAGCTGAGAGTGTGAGCTCCAGCTCCGCTCCGCTCATGCGCACTGAGTCTGGGAG 120
QY 1159 CAGTGCAGCTGAGAGCCGGAACCTGATCAGCAAGACCTCCAGAGGGGTATGAGAGTAA 1218
DB 121 CAGTGCAGCTGAGAGCCGGAACCTGATCAGCAAGACCTCCAGAGGGGTATGAGAGTAA 180
QY 1219 TCCTCCCACTGAGAGAGCTGCTGACTGCTCTCTGCTGCTGCTGCTGCTGCTGCT 1278
DB 181 TCCTCCCACTGAGAGAGCTGCTGACTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 1279 CCGTTTCAGAGCTGTAACCCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1338
DB 241 CCGTTTCAGAGCTGTAACCCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 1339 CCGCTGAGTGGGGGAGAGCCGAGTGGGAGAGCTGAGAGCCGCTGGGCTGGGG 1398
DB 301 CCGCTGAGTGGGGGAGAGCCGAGTGGGAGAGCTGAGAGCCGCTGGGCTGGGG 360
QY 1399 CCGCTGAGTGGGGGAGAGCCGAGTGGGAGAGCTGAGAGCCGCTGGGCTGGGG 1458
DB 361 CCGCTGAGTGGGGGAGAGCCGAGTGGGAGAGCTGAGAGCCGCTGGGCTGGGG 420
QY 1459 TGGCCGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1518
DB 421 TGGCCGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 480
QY 1519 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1578
DB 481 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 1579 CTGCAACAGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1638
DB 541 CTGCAACAGTATGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 600
QY 1639 GCGCTCTGAGAGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1698
DB 601 GCGCTCTGAGAGCTGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 660
QY 1699 CAGGCAACCCGAG 1758
DB 661 CAGGCAACCCGAG 720

QY 1759 ATACGCTGCTGAGAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1818
DB 721 ATACGCTGCTGAGAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 1819 GCGCCCTGAGAGCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878
DB 781 GCGCCCTGAGAGCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 1879 GATGAGAGTGTGAG 1938
DB 841 GATGAGAGTGTGAG 900
QY 1939 AGGCTGCTGAG 1998
DB 901 AGGCTGCTGAG 960
QY 1999 CTGCGGGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2058
DB 961 CTGCGGGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 2059 TGCAGCTGAG 2118
DB 1021 TGCAGCTGAG 1080
QY 2119 CCGGCTGAG 2178
DB 1081 CCGGCTGAG 1140
QY 2179 CAGCCGCTGAG 2238
DB 1141 CAGCCGCTGAG 1200
QY 2239 CAGCTCAACTTCCAG 2298
DB 1201 CAGCTCAACTTCCAG 1260
QY 2299 GCGCTGCTGAG 2358
DB 1261 GCGCTGCTGAG 1320
QY 2359 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2418
DB 1321 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 2419 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2478
DB 1381 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 2479 GAG 2538
DB 1441 GAG 1500
QY 2539 AGTGCAGTGTGAG 2598
DB 1501 AGTGCAGTGTGAG 1560
QY 2599 AGGAG 2658
DB 1561 AGGAG 1620
QY 2659 CTGAG 2718
DB 1621 CTGAG 1680
QY 2719 CAGGCTTCCCGAG 2778
DB 1681 CAGGCTTCCCGAG 1740
QY 2779 ATGTCGCGAG 2838
DB 1741 ATGTCGCGAG 1800

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp586J0619"
/tissue_type="uterus"

```

	gene	/clone_id="586 (synonym: hute1). Vector psportll; host DH10B; sites NotI + SalI/MluI"
CDS	1..1924	/dev_stage="adult"
	1..1924	/gene="DKEFp5860619"
	<1..1924	/gene="DKEFp5860619"
	/note="unknown"	
	/codon_start=2	
	/product="hypothetical protein"	
	/protein_id="CAB4378.1"	
	/db_xref="GI:4884140"	
	/db_xref="SPTREMBL:O9Y3W8"	
	/translation="TSSRSRLAFLAPPSATADASFPKCPVYVSLLIDEEP AGSGKAGGSLTRALPLPSSGLIVTMLMDPEVSSCGDLILFSRRKGQQA VSRPYRILLTFHOSMPTLHOCITRVLRQRDRDPASIDLIMACIHVRILWQ ROGRPPQKRRELVLAVQGPBLISVELIIATAEQLSDDEDAALCOARLILLS CGDDSEYKRYATEHLGGCILOMGDSVIGRRCDLILLYORREPLRVAYELHSR AASSVCCLDGLIHRTYLTLADTSRLLENNGDAMCARKIAYAHPLLHPML AALLRGRIINLEFEFOONHLSCTPIVGLILELTILELPYRSHOGAILDCLSFTL ILNRSSRHIAFINFWOAFIRYTHAKTYTNPAALISTDKRADPLHLSFMSDLMIK LIAGLSPSSDDPTBDGLDERESSAGSLPVSALTPLTAENAPMKRLSPQG LYAEILTSLSDISEMSRRRPILSFSTYLQIKMSIEDCCNIALSIARSQNSP IAAALTLPWICLGSDQDFEVOTALNLPLEVALLCQSHAVALHRAFLVANGTGNDP AQSIALRILIMEAW"	
	2224	
polya_site	383 a	754 c 712 g 392 t
BASE COUNT		
ORIGIN		
Query Match	64.1%	Score 2203.8; DB 9; Length 2241;
Best Local Similarity	99.9%;	Pred. No. 1.6e-304;
Matches 2205; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
Oy	1231	GAGGAGCTGCAGTCAGCATTTCTTCTGCCACTCGGANTCTGCTCCCGGTTCACCC 1290
Dd	20	GCGAGCTGCTGACATTCATTCTCTGCCACTCGGANTCTGCTCCCGGTTCACCC 79
Oy	1291	TGTAAAGCCGTTTGATGATGAGCTCCCGCTGCTGTGAAGAAGAGAGCCCCCTGGCTGG 1350
Dd	80	TGTAAAGCCGTTTGATGATGAGCTCCCGCTGCTGTGAAGAAGAGAGCCCCCTGGCTGG 139
Oy	1351	GGGAAGCCGGGATGTCGACGGTGGCACTGGAGGCGCTCGGCTGCGGACCCTTCGTCAGCC 1410
Dd	140	GGGAAGCCGGGATGTCGACGGTGGCACTGGAGGCGCTCGGCTGCGGACCCTTCGTCAGCC 199
Oy	1411	CTCCAGTAGGACATGGCTGGAAGAAGCTGGAACCCCGAGGTGGTGAACAAGTCTCCCGCACCTG 1470
Dd	200	CTCTAATGATGACATGGCTGGAAGAAGCTGGAACCCCGAGGTGGTGAACAAGTCTCCCGCACCTG 259
Oy	1471	CAGCTCAGAGCTGCTTCTTCTCCCGAGAGAAGGGCAAAGGTCAAGGCCCAAGTGGGCCCTCGTTC 1530
Dd	260	CAGCTCAGAGCTGCTTCTTCTCCCGAGAGAAGGGCAAAGGTCAAGGCCCAAGTGGGCCCTCGTTC 319
Oy	1531	CGTCCCTAACCTCTCTAACCCCTTTCACAGCATCAAGTCCAGCTGGCCCAACATGACACCAAGTGC 1590
Dd	320	CGTCCCTAACCTCTCTAACCCCTTTCACAGCATCAAGTCCAGCTGGCCCAACATGACACCAAGTGC 379
Oy	1591	ATTCGAGTCTCTGGGCAAGAGACCGGGAAATAAGAGTTGACCCCTCTGCTCTTGAGAC 1650
Dd	380	ATTCGAGTCTCTGGGCAAGAGACCGGGAAATAAGAGTTGACCCCTCTGCTCTTGAGAC 439
Oy	1651	TTCCTCTGGGACCTGCATCATATGTTCTCTGCAATGTGGACAGGGGCGGGAACCAAGGGACCCCG 1710
Dd	440	TTCCTCTGGGACCTGCATCATATGTTCTCTGCAATGTGGACAGGGGCGGGAACCAAGGGACCCCG 499
Oy	1711	CAGAAGCGGCGGAGAGAGCTGGTGTCTCGAGGGCCCGAGACTCATAGACTTGTTG 1770
Dd	500	CAGAAGCGGCGGAGAGAGCTGGTGTCTCGAGGGCCCGAGACTCATAGACTTGTTG 559
Oy	1771	GAGGTATCTTGGGCGAGAGGGGAGACCGGGAGATCAAGACAGGAGACCAAGCTCGCTGACAC 1830
Dd	560	GAGGTATCTTGGGCGAGAGGGGAGACCGGGAGATCAAGACAGGAGACCAAGCTCGCTGACAC 619

[illegible]

SOURCE house mouse,
ORGANISM *Mus musculus*
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; *Mus* 1 (bases 1 to 2730)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2001) National Institutes of Health

tissue procurement: Gilbert Smith, Ph.D.
 cDNA library preparation: life technologies, inc.
 cDNA library arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amge@bcm.tmc.edu
 Guaratine, P.H., garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

The clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Source

Location/Qualifiers

1. 2730

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="rvb/a"

/clone="IMAGE:3601115"

/tissue-type="Mammary tumor, Metastatic; human; alpha model, 10 month old virgin mouse. Taken by biopsy."

/clone_id="NCI_GAP_Mam1"

/lab_host="DH10B"

/note="Vector: pCMV-Sport6"

<1. 2458

/codon_start=2

/product="Unknown (protein for IMAGE:3601115)"

/protein_id="AAH10333.1"

/db_xref="GI:16307574"

/translation="TALALQALQELARVQGNPEVDITVRLQAMTLLSPHG
TALAMHSHFSLQPLMLQYORAVPDTPGSLFLKMOILQDLSPPVGP
QALKEPATRYSAHRISDVSGLHLADLSPHGLVSTRAVATLINSKCP
VEPELISKVRLGLIEYSPHEILLTALSTMTSCPSGPIVVSLLQKEE
LLGSKQEVGASTAMRLGASGLVDHETLDEPVSCPDLMWKLFSRKHG
LSAOLVSPRYLLALTRQASSTLHLCILVILGSRQRLDPSALDILFMRH
IMQGRDPTQKRRRLVHVOGPELILVILSEATSDQDPSAALTLQETPL
ILSCRSNDSEIKVLEHLETCIQMDSVLQQRDLILQYORAVRYVPEVIL
QSEATSSSICKLDGIVHETLADTSDRSERSEVADAMCRILAVHVPVIL
LPMIALHGRTHINFEERFROOHLAFELVIGILLQPRVQSHQALMDLCKSF
RLILNTRKSRHIAPIFSKFPVIRHYVCSAPYATLZKNAEPLHDSNDIV
MKSILAGLIDPSRQDGLDEBDEBSAGSPVLSPLTVADVAPHKRL
SRGAVDYLETSPIDMSRRRVEYFSTNQLRMSAESRCHLAFSLALSL
NPSIADPLTFMTCLSDNFVEVYQALRNLPETVILCOEHAVALHRAFLVYTGQ
IDTSAQISELKLHMEAYM"

BASE COUNT 541 a 877 c 786 g 526 t
ORIGIN

Query Match 50.2%; Score 1726; DB 10; Length 2730;
Best Local Similarity 82.0%; Pred. No. 1.7e-236;
Matches 2016; Conservative 0; Mismatches 435; Indels 9; Gaps 2;

687 CTGTGCTCTGCGCCCTGACAGAGCCCTGAGGACAGAGAGCTGAGGCTGCTCCAGAGGCAAG 746
1 CCTTGGCCCTGGACACTGACAGCAACCTGAGGCAAGAGTTGGCTGTGCTCCGCAAGGGA 60
747 CCCCAGAGTCCCGGAGCATCAGGTGCTGCTCTGACAGGCTTGGCCACCTGCTCAAGCTC 806
61 CCCCAGAGTCCCGGAGCATCAGGTGCTGCTCTGACAGGCTTGGCCACCTGCTCAAGCTC 120
807 CCCCAGAGTCCCGGAGCATCAGGTGCTGCTCTGACAGGCTTGGCCACCTGCTCAAGCTC 866
121 CCCCAGAGTCCCGGAGCATCAGGTGCTGCTCTGACAGGCTTGGCCACCTGCTCAAGCTC 180
867 GCTGAGCAGCTGCTGACAGGCTGCTGCTGACAGGCTTGGCCACCTGCTCAAGCTC 926
181 GATGCGCAGCTGCTGACAGGCTGCTGCTGACAGGCTTGGCCACCTGCTCAAGCTC 240
927 CTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
241 CTTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986
301 GCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
987 GCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1046
301 GCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
1047 TGAATGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
361 GATGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
1107 GATGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1166
421 GATGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
1167 CAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1226

481 TGTGAGCTGAGCTGATCATCAAAAGCTCCCGGGGCTCATGAGATGCTGCTCCCGCA 540
1227 CCGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286
541 CTTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
1287 AGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1343
601 AGCTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
1344 GCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1403
661 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
1404 GTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1463
721 TTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
1464 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1517
781 TGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
1518 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1577
841 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
1578 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1637
901 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
1638 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1697
961 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1720
1698 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1757
1021 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
1758 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1817
1081 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1840
1818 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1877
1141 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1900
1878 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1937
1201 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1960
1938 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1997
1261 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2057
1998 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2120
1321 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2157
2058 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2180
1381 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2217
2118 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2237
1441 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
2178 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2237
1501 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
2238 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2297

Db 1561 TCATCTGAACTCCAGAGATCCGAGACAGAACACCTGCTCTCTGATGATGCT 1620

Qy 2298 GGGGCTGCTGAGATGCTGAGAGCGAGCTGTTCCGAGGACACACAGGGGGGCTG 2357

Db 1621 GGGGATCTGAGAGCTGCTGAGAGCGAGCTGTTCCGAGGAGAACACAGGGGGGCTG 1680

Qy 2358 GGAATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2417

Db 1681 GGAATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1740

Qy 2418 GGGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 2477

Db 1741 GGGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1800

Qy 2478 AGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2537

Db 1801 AGCAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860

Qy 2538 CAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2597

Db 1861 CAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920

Qy 2598 CAGGAGCCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2657

Db 1921 GAGATGAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980

Qy 2658 CAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2717

Db 1981 CAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040

Qy 2718 AGCGCTTCCCGGGGGCCAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2777

Db 2041 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100

Qy 2778 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2837

Db 2101 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160

Qy 2838 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2897

Db 2161 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220

Qy 2898 GAGCAGCCAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2957

Db 2221 GAGCAGCCAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280

Qy 2958 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3017

Db 2281 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340

Qy 3018 AGAGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3077

Db 2341 AGAGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400

Qy 3078 CCGCAGGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3137

Db 2401 CCGCAGGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460

RESULT 8
HSM802043
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFp434C0126 (from clone DKFp434C0126);
partial cds.
ACCESSION AL137358
VERSION AL137358.1 GI:8607876
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1141)
Ansojge,W., Winkler,U., Meves,H.W., Gassenhuber,U. and Wiemann,S.

TITLE Direct Submission
JOURNAL Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFp434C0126) is available at the RZPD in Berlin.
Berlin-Charlottenburg, Germany; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

source
1..1141
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFp434C0126"
/issue="testis"
/clone_id="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..934
/gene="DKFp434C0126"
1..934
/gene="DKFp434C0126"
/note="DKFp434C0126, partial"
/codon_start=2
/product="hypothetical protein"
/protein_id="CA870710.1"
/db_xref="GI:6807877"
/translation="LLOPHVRSRHOGALMDLSTFRLILNKRSSRLAFLNFKY
DEEERSSAGSLPIYSVLETPITAEAPYMKLSGQIVDLLEVLSDIDESSR
RPELISFTSNDRMLSSAEBCCKNAFLARSQNSPSIAAFLPIPMCLISDF
EVDVIRKNIPEKALLCGEHAVALHRAFGGCHVDPQADRALRPEDEPAVGSDVS
LMDPTQAPARVPEDPRKAKRAVALIVCRR"
1126

BASE COUNT 209 a 384 c 339 g 209 t
ORIGIN

Query Match 31.4%; Score 1079.4; DB 9; Length 1141;
Best Local Similarity 98.4%; Pred. No. 2.3e-144;
Matches 1111; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

Qy 2310 GCTGCTGAGCGGAGCGAGTTCGCGAGGAGACACAGGGGGGCTGCTGCTGCT 2369

Db 1 GCTGCTGAGCGGAGCGAGTTCGCGAGGAGACACAGGGGGGCTGCTGCTGCTGCT 60

Qy 2370 GTCCCTCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2429

Db 61 GTCCCTCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

Qy 2430 CAACAGTTTGAGCAATTCATCAATTAATTAATTAATTAATTAATTAATTAATTA 2489

Db 121 CAACAGTTTGAGCAATTCATCAATTAATTAATTAATTAATTAATTAATTAATTA 180

Qy 2490 CTTCCTGAGAGAGAGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2549

Db 181 CTTCCTGAGAGAGAGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

Qy 2550 GATGCTGAATTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2609

Db 241 GATGCTGAATTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

Qy 2610 AGGCGGAG 2669

Db 301 AGGCGGAG 360

Qy 2670 CTCCTGTTACCCCTCTGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2729

Db 361 CTCCTGTTACCCCTCTGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

QY 2730 GGGCCAAAGGCGGAGGATCTGCTGAGGATCTGAGTACATAGAGAGATGTCCGGCG 2789
 Db 421 GGGCCAAAGGCGGAGGATCTGCTGAGGATCTGAGTACATAGAGAGATGTCCGGCG 480
 QY 2790 GAGACCCGAGATCTGAGCTCTCTGACCAACCTGACGCGGCTGATGAGCTCGGCGA 2849
 Db 481 GAGACCCGAGATCTGAGCTCTCTGACCAACCTGACGCGGCTGATGAGCTCGGCGA 540
 QY 2850 GAGATGTTGCCGCAACCTGCTGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2909
 Db 541 GAGATGTTGCCGCAACCTGCTGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 600
 QY 2910 CATTCGACCGCTTCTCTCCAGCTGACATGATGATGATGATGATGATGATGATGAT 2969
 Db 601 CATTCGACCGCTTCTCTCCAGCTGACATGATGATGATGATGATGATGATGATGAT 660
 QY 2970 GGTGACGACGCGCTTCCGGAACCTGAGTACCTGCTCTCTGCTGCTGCTGCTGCT 3029
 Db 661 GGTGACGACGCGCTTCCGGAACCTGAGTACCTGCTCTCTGCTGCTGCTGCTGCT 720
 QY 3030 TGTGCTGCTCAGCGGCGCC-TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3088
 Db 721 TGTGCTGCTCAGCGGCGCC-TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 3089 AGATCTCCGAGCGCCGAGGATCTGATGATGATGATGATGATGATGATGATGATGAT 3148
 Db 781 AGATCTCCGAGCGCCGAGGATCTGATGATGATGATGATGATGATGATGATGATGAT 840
 QY 3149 GACCCCGCTCAGCGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3208
 Db 841 GACCCCGCTCAGCGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 3209 AGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3268
 Db 901 AGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 3269 GCCCAGGAGCAATATCTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3328
 Db 961 GCCCAGGAGCAATATCTCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 3329 CGTCCAGCAAGCGCTCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3388
 Db 1021 CGTCCAGCAAGCGCTCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3448
 QY 3389 CCGGCGCAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 Db 1081 CCGGCGCAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128

RESULT 9
 AC102953/c 148996 bp DNA linear PRI 01-JUN-2002
 LOCUS Homo sapiens BAC clone RP11-1246C19 from 7, complete sequence.
 DEFINITION AC102953
 VERSION AC102953.5 GI:21307577
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 148996)
 AUTHOR(S) Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 9847074
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 148996)
 AUTHOR(S) Du, H. and Kozlovicz, A.
 TITLE The sequence of Homo sapiens BAC clone RP11-1246C19
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 148996)
 AUTHOR(S) Waterston, R.H.
 TITLE Direct Submission

JOURNAL

Submitted (23-NOV-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 148996)
 AUTHOR(S) Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (22-FEB-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 148996)
 AUTHOR(S) Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAR-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 6 (bases 1 to 148996)
 AUTHOR(S) Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 7 (bases 1 to 148996)
 AUTHOR(S) Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jun 1, 2002 this sequence version replaced gi:15974660.

REFERENCE

Center: Washington University Genome Sequencing Center
 Center code: WUSGC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: saplens@wustl.wustl.edu
 Summary Statistics
 Center project name: H_NHL246C19

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NIGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
<http://www.nih.gov/DIR/CTB/CHR7>, send
mailto:egreen@nigri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
 Tatem, K., Catalanese, J.J., and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Piefer de Jong
 and coworkers at <http://www.chori.org>
 VECTOR: pBRac3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-369D24, 2000 bp overlap;
 the clone sequenced to the right is RP11-16P10, 2000 bp overlap.
 Actual start of this clone is at base position 174583 of
 RP11-369D24; actual end is at base position 63205 of RP11-16P10.

repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="7"
repeat_region	/map="7"
repeat_region	/clone="RP11-1246C19"
repeat_region	/clone_id="RPCL-11"
repeat_region	585..893
repeat_region	/rpt_family="Alu"
repeat_region	2518..2618
repeat_region	/rpt_family="MIR"
repeat_region	2708..2769
repeat_region	/rpt_family="MIR"
repeat_region	2873..3034
repeat_region	/rpt_family="MIR"
repeat_region	3347..3438
repeat_region	/rpt_family="L1"
repeat_region	3439..3574
repeat_region	/rpt_family="Alu"
repeat_region	3575..3857
repeat_region	/rpt_family="Alu"
repeat_region	3858..4034
repeat_region	/rpt_family="Alu"
repeat_region	4035..4628
repeat_region	/rpt_family="L1"
repeat_region	4629..4714
repeat_region	/rpt_family="Alu"
repeat_region	4796..5037
repeat_region	/rpt_family="Alu"
repeat_region	5039..5286
repeat_region	/rpt_family="L1"
repeat_region	6609..6933
repeat_region	/rpt_family="Alu"
repeat_region	6938..7061
repeat_region	/rpt_family="L1"
repeat_region	7145..7243
repeat_region	/rpt_family="MERL_type"
repeat_region	7256..7559
repeat_region	/rpt_family="Alu"
repeat_region	9418..9677
repeat_region	/rpt_family="L2"
repeat_region	9758..9876
repeat_region	/rpt_family="MERL_type"
repeat_region	9886..9926
repeat_region	/rpt_family="TTCCn"
repeat_region	9927..10207
repeat_region	/rpt_family="Alu"
repeat_region	10208..10220
repeat_region	/rpt_family="MERL_type"
repeat_region	10227..10366
repeat_region	/rpt_family="L2"
repeat_region	11484..11545
repeat_region	/rpt_family="MIR"
repeat_region	12151..12259
repeat_region	/rpt_family="MIR"
repeat_region	12792..13111
repeat_region	/rpt_family="Alu"
repeat_region	14332..14451
repeat_region	/rpt_family="MIR"
repeat_region	14794..14822
repeat_region	/rpt_family="AT-rich"
repeat_region	14828..15133
repeat_region	/rpt_family="Alu"
repeat_region	15194..15588
repeat_region	/rpt_family="MIR"
repeat_region	16146..16439
repeat_region	/rpt_family="Alu"

```
/rpt_family="(CACCC)n  
23254. .23354  
/rpt_family="MIR"
```

[illegible]

KEYWORDS
SOURCE
ORGANISM

HRG: HRGS_PHASE1.
Rattus norvegicus.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 144097)

Mazy,D.M., Adams,C., Adio-Ochola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbieri,J., Benton,T., Bimaga,K., Blankenburg,K., Bonin,D.,
Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinn,H.E.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Ehrhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabali,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,T., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,X., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mathiney,E., McLeod,M.P., Meador,M., Mel,G., Melker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogun,M., Okunnu,G.,
Ortugue,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Roite,M., Ruiz,S., Savery,G.,
Schaefer,S., Scott,G., Shen,H., Shoostari,N., Slason,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Uman,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE

Unpublished

JOURNAL

2 (bases 1 to 144097)

REFERENCE

Worley,K.C.

TITLE

Direct Submission

JOURNAL

Submitted (05-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 144097)

TITLE

Worley,K.C.

JOURNAL

Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jul 18, 2002 this sequence version replaced gi.20452937.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GWMB

Center clone name: CH230-101110

----- Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.990329
Consensus quality: 62007 bases at least Q40
Consensus quality: 68326 bases at least Q30
Consensus quality: 73396 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 72 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1060:	contig of 1060 bp in length
1061	1160:	gap of unknown length
1161	2270:	contig of 1110 bp in length
1271	2370:	gap of unknown length
2371	3371:	contig of 1001 bp in length
3372	3471:	gap of unknown length
3472	4834:	contig of 1363 bp in length
4835	4935:	gap of unknown length
4935	5992:	contig of 1058 bp in length
5993	6092:	gap of unknown length
6093	7670:	contig of 1578 bp in length
7671	7770:	gap of unknown length
7771	9312:	contig of 1442 bp in length
9313	9312:	gap of unknown length
9313	10541:	contig of 1229 bp in length
10542	10641:	gap of unknown length
10642	12072:	contig of 1431 bp in length
12073	12172:	gap of unknown length
12173	13552:	contig of 1380 bp in length
13553	13652:	gap of unknown length
13653	15096:	contig of 1444 bp in length
15097	15196:	gap of unknown length
15197	16222:	contig of 1026 bp in length
16223	16322:	gap of unknown length
16323	17655:	contig of 1333 bp in length
17656	17755:	gap of unknown length
17756	18939:	contig of 1184 bp in length
18940	19039:	gap of unknown length
19040	20566:	contig of 1527 bp in length
20567	20666:	gap of unknown length
20667	21987:	contig of 1321 bp in length
21988	22087:	gap of unknown length
22088	22119:	contig of 1032 bp in length
22120	23219:	gap of unknown length
23220	24254:	contig of 1035 bp in length
24255	24354:	gap of unknown length
24355	25397:	contig of 1043 bp in length
25398	25497:	gap of unknown length
25498	27045:	contig of 1548 bp in length
27046	27145:	gap of unknown length
27146	28578:	contig of 1433 bp in length
28579	28678:	gap of unknown length
28679	29972:	contig of 1294 bp in length
29973	30072:	gap of unknown length
30073	31567:	contig of 1495 bp in length
31568	31667:	gap of unknown length
31668	33101:	contig of 1434 bp in length
33102	33201:	gap of unknown length
33202	34444:	contig of 1243 bp in length
34445	34544:	gap of unknown length
34545	35824:	contig of 1280 bp in length
35826	35924:	gap of unknown length
35925	37318:	contig of 1394 bp in length
37319	37418:	gap of unknown length
37419	38684:	contig of 1266 bp in length
38685	38784:	gap of unknown length
38785	40419:	contig of 1635 bp in length
40420	40519:	gap of unknown length

```

* 40520 42463: contig of 1944 bp in length
* 42464 42563: gap of unknown length
* 42564 44280: contig of 1717 bp in length
* 44281 44380: gap of unknown length
* 44381 45965: contig of 1385 bp in length
* 45966 46065: gap of unknown length
* 46066 47557: contig of 1392 bp in length
* 47558 47557: gap of unknown length
* 47558 49418: contig of 1861 bp in length
* 49419 49518: gap of unknown length
* 49519 50846: contig of 1328 bp in length
* 50847 50946: gap of unknown length
* 50947 53079: contig of 2133 bp in length
* 53080 53179: gap of unknown length
* 53180 54796: contig of 1617 bp in length
* 54797 54896: gap of unknown length
* 54897 55980: contig of 1084 bp in length
* 55981 56080: gap of unknown length
* 56081 58603: contig of 2522 bp in length
* 58603 58702: gap of unknown length
* 58703 60423: contig of 1721 bp in length
* 60424 62328: gap of unknown length
* 62329 62428: contig of 1805 bp in length
* 62429 64486: gap of unknown length
* 64487 64586: gap of unknown length
* 64587 66593: contig of 2073 bp in length
* 66593 66759: gap of unknown length
* 66760 68152: contig of 1393 bp in length
* 68153 68252: gap of unknown length
* 68253 70718: contig of 2466 bp in length
* 70719 70818: gap of unknown length
* 70819 72273: contig of 1455 bp in length
* 72274 72373: gap of unknown length
* 72374 74454: contig of 2081 bp in length
* 74455 74555: gap of unknown length
* 74556 76795: contig of 2241 bp in length
* 76796 76895: gap of unknown length
* 76896 79006: contig of 2111 bp in length
* 79007 79107: gap of unknown length
* 79108 82084: contig of 2878 bp in length
* 82085 82184: gap of unknown length
* 82185 83488: contig of 1305 bp in length
* 83489 83588: gap of unknown length
* 83590 86730: contig of 3141 bp in length
* 86731 86830: gap of unknown length
* 86831 88726: contig of 1896 bp in length
* 88727 88827: gap of unknown length
* 88828 90953: contig of 2127 bp in length

```

Query Match 5.78; Score 194.4; DB 2; Length 144097;
 Best Local Similarity 84.28; Pred. No. 9.3e-19;
 Matches 219; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```

OY 632 TCCAGATTTTCCCGTCAGACCGCGGACCGGTCGAGACTTCACATTCCTCCGCGCCCGCGG 691
DB 51563 TCCAGATTTTCCCGTCAGACCGCGGACCGGTCGAGACTTCACATTCCTCCGCGCCCGCGG 51504
OY 692 CCTCGCCCTTCAGACGAGCCCTGGGCGACGAGCTGCGCCGCGTCGTCGACGAGCGCGCG 751
DB 51503 CCTCGCCCTTCAGACGAGCCCTGGGCGACGAGCTGCGCCGCGTCGTCGACGAGCGCGCG 51444
OY 752 AGGTGCGCGGATACAGGTGCTGCTTCGAGCGCCCTCGGCAACCTGCTAGCTCCGAC 811
DB 51443 AGGTGCGCGGATACAGGTGCTGCTTCGAGCGCCCTCGGCAACCTGCTAGCTCCGAC 51384
OY 812 ACGGCGGTGCGCTGATGATGTCATGACAGCGTACGACATTCCTGAGCTGCCCGCTCTGCG 871
DB 51383 ACGGCGGTGCGCTGATGATGTCATGACAGCGTACGACATTCCTGAGCTGCCCGCTCTGCG 51324
OY 872 GCCAGCTCTGCCAGTACAG 891
DB 51323 GCCAGCTCTGCCAGTACAG 51304

```

FEATURES

Source

misc_feature

misc_feature

misc_feature

misc_feature

RESULT 11
 AC130221
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@wustl.wustl.edu
 Project Information
 Center project name: M.BA0168E11
 Summary Statistics
 Sequencing vector: M13; 08
 Chemistry: Dye-primer ET; 08 of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 217436 bases at least Q40
 Consensus quality: 217966 bases at least Q30
 Consensus quality: 218400 bases at least Q20
 NOTE: This is a 'working draft' sequence. It currently
 consists of 6 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1
 2186: contig of 2186 bp in length
 2187: gap of unknown length
 2287: contig of 4032 bp in length
 6318: gap of unknown length
 6319: gap of unknown length
 6418: gap of unknown length
 33825: contig of 27407 bp in length
 33826: gap of unknown length
 33926: contig of 46123 bp in length
 80048: gap of unknown length
 80049: gap of unknown length
 80149: gap of unknown length
 128323: contig of 48175 bp in length
 128324: gap of unknown length
 219964: contig of 91541 bp in length.
 Location/Qualifiers
 1. 219964
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="YMK"
 /clone="RP23-168E11"
 1. 2186
 /note="assembly_name:Contig13"
 2287. 6318
 /note="assembly_name:Contig14"
 6419. 33825
 /note="assembly_name:Contig15"

TITLE Direct Submission
JOURNAL Submitted (01-MAY-1998) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road

Ager, S., Al, R., Baldwin, D., Banzon, J., Beeson, K. Y., Busan, D. A., Carlson, J. W., Center, A., Champagne, M., Daveyport, L. B., Dietz, S. M., Dodson, K., Dorsett, V., Doum, L. E., Doyle, C., Drenek, D., Farfan, D., Ferreira, S., Frise, E., Galle, R. F., Gary, N. S., George, R. A., Gonzalez, M., Hunk, T., Inertea, A., Johnson, P.

db 133048 AAGACTTAATTGGAGGAGCTCTCAAGACTCAAGCCGAGATCGTCTCTTTGTGC 133107
Y 308 AGTCGTTGGCAGACCCGCTGTCCAGCATGAGCAAGACCTCCAGTCTCTGGACACGGAG 367
db 133108 AAACTTTGGCAAGCCGGTCAATCGATGATCCAAAGCTGTGGCAATGCTTGGAACCGTG 133167
Y 368 TGGCCACGAGACCCCGACGACTGTGAGCAGAAACATCATGAGCAAGATTAAAGGCCCCACC 427
db 133168 TACTCGAGCAGTTTATCTGTTGAAGATGCCATTCTGAAACAAGGCCCTATCTAGGCCACG 133227

DEFINITION	303523 dp	DNA	linear	INV 28-JUN-2002
	<i>Drosophila melanogaster</i> 2R	section 70 of 74 of the complete arm,		
ACCESSION	AE003462	AE003599	AE002575	
VERSION	AE003462.2	GI:21626662		

KEYWORDS

fruit fly

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 303823)

Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Wei,K.H., Doyle,C., Baxter,B.G., Helt,G., Chapple,M., Pfeiffer,B.D., Miklos,G.L., Abril,J.F., Abmayyan,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolintsov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C., Busan,D.A., Butler,H., Cadden,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieres,S., Fleischmann,W., Folsler,C., Gabrielian,A.E., Garg,N.S., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wel,M.H., Ibegwam,C., Jalili,M., Kalush,F., Karpen,G.H., Ke,Z., Kennedy,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,X., Lin,X., Lin,X., Matrei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Mekhalov,G., Milshina,N.V., Mobarry,C., Morris,J.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nizny,D.M., Nelson,D.L., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muskern,D.R., Paclet,J.M., Palazolo,M., Pittman,G.S., Pan,S., Pollard,J.D., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kimmos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sud,E., Szitskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yen,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,X., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

TITLE

The genome sequence of *Drosophila melanogaster*

Science 287 (5461), 2185-2195 (2000)

JOURNAL

MEDLINE

20196006

10731132

2 (bases 1 to 303823)

Amanatides,P.G., Brandon,R.C., Rogers,Y., An,H., Baldwin,D., Barzon,J., Beeson,K.Y., Busan,D.A., Carlson,J.W., Center,A., Champ,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferrieres,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalili,M., Kruse,D., Li,P., Matrei,B., Moshrefi,A., McIntosh,T.C., Patayes,V., Murphy,B., Nelson,C., Nelson,K.A., Nungo,J., Paclet,J., Pargacs,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Stapleton,M., Strong,R., Szitskas,R., Tector,C., Tyler,D., Williams,S.M., Young,J.S., Svitskas,R., Tector,C., Tyler,D., Williams,S.M., Young,J.S., Smith,H.O., Venter,J.C. and Rubin,G.M.

TITLE

Sequencing of *Drosophila melanogaster* genome

Unpublished

3 (bases 1 to 303823)

Campbell,K., Crosby,M.A., Matthews,B.B., Bayraktaroglu,L., Smith,C.D., Tupy,J.L., Bergman,C.M., Berman,B.P., Carlson,J.W., Celniker,S.E., Clamp,M.E., Drysdale,R.A., Emmert,D., Frise,E., de Grey,A.D.N.J., Harris,M.N., Krommler,B., Marshall,B., Millburn,G.H., Richter,J., Russo,S., Searle,S.M.J., Smith,E., Smu,S., Smitnick,F., Whitfield,E.J., Yamada,C., Ashburner,M., Gelbart,W.M., Rubin,G.M., Mungall,C.J. and Lewis,S.E.

TITLE

Annotation of *Drosophila melanogaster* genome

Unpublished

4 (bases 1 to 303823)

Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,J.C.

REFERENCE

1 (bases 1 to 303823)

Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

JOURNAL

MEDLINE

20196006

10731132

2 (bases 1 to 303823)

Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

COMMENT

Submitted (31-MAR-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA

FEATURES

On Jun 28, 2002 this sequence version replaced gi:7291637, location/Qualifiers

source

1. 303823

organism

/organism="Drosophila melanogaster"

taxonomy

/db_xref="taxon:7227"

chromosome

/chromosome="2"

map

/map="2R"

complement

/complement(11873..23611)

map

/map="60A9-60A10"

db_xref

/db_xref="FLYBASE:FBan0011169"

db_xref

/db_xref="FLYBASE:FBgn0034905"

complement

/complement(join(11873..12379,18020..18091))

note

/note="CG11169 gene product"

codon_start

/codon_start=1

product

/product="CG11169-PA"

protein_id

/protein_id="AA047060.1"

db_xref

/db_xref="FLYBASE:FBan0011169"

db_xref

/db_xref="FLYBASE:FBgn0034905"

translation

/translation="MSSPVEETALLVSSYTCQGLPDLKMPEDQSVFVHTD SGNNLQEPYSAIKSHNVSFARQHLNLSLATVGLFKAPQLMMVTGSEPMAT HEDLQIVHVOYSRVYSIDLKRYKRYV"

complement

/complement(16742..17216)

gene

/gene="CG15800"

map

/map="60A10-60A10"

db_xref

/db_xref="FLYBASE:FBan0015800"

db_xref

/db_xref="FLYBASE:FBgn0034904"

complement

/complement(16742..17216)

note

/note="CG15800 gene product"

codon_start

/codon_start=1

product

/product="CG15800-PA"

protein_id

/protein_id="AA047061.1"

db_xref

/db_xref="FLYBASE:FBan0015800"

db_xref

/db_xref="FLYBASE:FBgn0034904"

translation

/translation="MVFETEDGRTISVALLAKRSIVYREMRGQVDEEDVYDL HGISNVLLKILMAVEHEKEEPAWDSKDLPAETELQISDMKFEFLREIDSIC KIMEGCVLDIPMLYKCAQKIVFLSSRPEIQRSGTVEKIPFEDHLEIOIRE"

complement

/complement(16742..17216)

gene

/gene="CG30177"

map

/map="60A10-60A10"

db_xref

/db_xref="FLYBASE:FBan0030177"

db_xref

/db_xref="FLYBASE:FBgn0050177"

complement

/complement(16742..17216)

note

/note="CG30177-PA"

map

/map="60A10-60A10"

db_xref

/db_xref="FLYBASE:FBan0030177"

db_xref

/db_xref="FLYBASE:FBgn0050177"

complement

/complement(16742..17216)

gene

/gene="CG30177-PA"

